

*Andrzej Bochniak, Mirosława Wesółowska-Janczarek  
Katedra Zastosowań Matematyki  
Akademia Rolnicza w Lublinie*

## **PROBLEM DOBORU WSPÓLNEJ KRZYWEJ DLA DWÓCH REPLIKACJI NA PRZYKŁADZIE PROCESU KIEŁKOWANIA ZIAREN ZBÓŻ STYMULOWANYCH POLEM MAGNETYCZNYM**

### **Streszczenie**

W pracy przedstawiony jest problem doboru wspólnej krzywej obrazującej przebieg badanego procesu, gdy mamy do dyspozycji dwie replikacje. Wybrana krzywa powinna być dobrym dopasowaniem dla obu replikacji. Do analizy, na podstawie której zostanie przedstawiony problem dopasowania wspólnej krzywej, wybrano krzywą logistyczną, przy pomocy której daje się oszacować przebieg procesu kiełkowania ziaren zbóż stymulowanych polem magnetycznym. Jako kryterium dobroci dopasowania krzywej zastosowano, często w tym celu wykorzystywany, współczynnik determinacji  $R^2$ . Podano sposób wstępnej oceny, czy do otrzymanych replikacji daje się dopasować wspólną krzywą logistyczną, która może być oszacowaniem zarówno dla pierwszej, jak i drugiej replikacji. Praca ma na celu przedstawienie różnych możliwości doboru takiej krzywej z wykorzystaniem współczynnika  $R^2$  jako kryterium dobroci dopasowania krzywej i zwrócenie uwagi na zaobserwowane własności tych metod.

**Słowa kluczowe:** dopasowanie krzywej, krzywa logistyczna, współczynnik  $R^2$ , stymulacja polem magnetycznym

### **Wstęp i cel pracy**

Procesy zachodzące w czasie opisuje się przez dopasowywanie krzywej do otrzymanych w wyniku doświadczenia danych liczbowych. Istnieją różne sposoby oceniania dobroci jej dopasowania. Często stosowanym kryterium jest współczynnik determinacji  $R^2$ . W literaturze znane są metody dopasowania krzywych dla ciągu pojedynczych obserwacji, jak też i dopasowania wspólnej krzywej przy większej liczbie powtórzeń.

Celem pracy jest przedstawienie różnych możliwości doboru wspólnej krzywej obrazującej badany proces. Zwrócono także uwagę na zaobserwowane własności tych metod. Rozważania przedstawione zostały na przykładzie procesu kiełkowania ziaren zbóż poddanych przedświejnej stymulacji polem magnetycznym, z wykorzystaniem współczynnika  $R^2$  jako kryterium dobroci dopasowania krzywej. Podczas próby dopasowania wspólnej krzywej dla kilku replikacji wynikły pewne problemy, które zostały przybliżone w pracy.

### **Materiał doświadczalny**

Do badań użyto nasion pochodzących z Instytutu Genetyki i Hodowli Roślin Akademii Rolniczej w Lublinie. Były to trzy formy zbóż: pszenżyto „Presto”, pszenperz PPG 115 oraz mieszaniec „Presto” x PPG 115. Partie nasion wszystkich form poddano działaniu zmiennego pola magnetycznego o indukcji 30 mT w dwóch czasach ekspozycji 8 i 30 sekund zgodnie z przyjętą metodyką badań [Pietruszewski 1998]. W celach porównawczych dla poszczególnych odmian pobrano nasiona kontrolne. W Katedrze Fizyki Akademii Rolniczej w Lublinie zostały przeprowadzone badania mające na celu określenie procesu kiełkowania poszczególnych form zbóż i wpływu na ten proces stymulacji materiału siewnego. Na każdym poletku, na którym badano proces kiełkowania, wysiano po 500 ziaren.

### **Model matematyczny procesu kiełkowania ziaren zbóż**

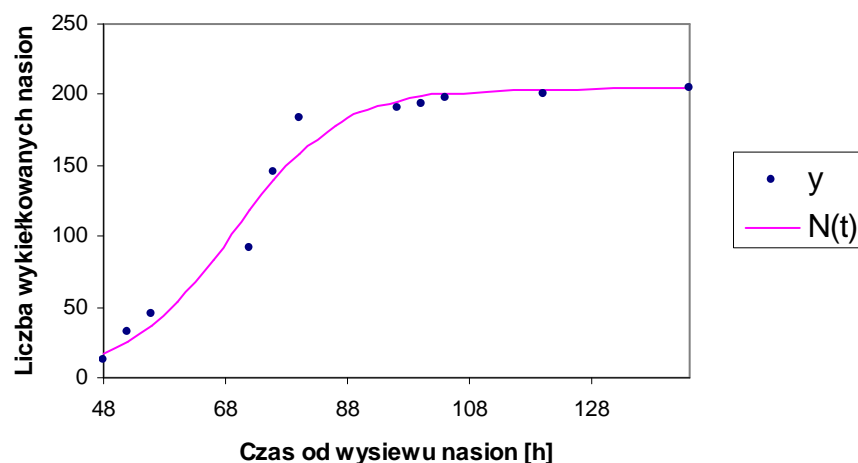
Proces kiełkowania nasion poddanych przedświejnej stymulacji polem magnetycznym można opisać przy pomocy modelu matematycznego, który wyrażony jest funkcją logistyczną [Pietruszewski i Kornarzyński 1999, 2001] postaci:

$$N(t) = \frac{N_k}{1 + (N_k - 1)e^{-\alpha N_k (t - t_0)}} \quad (1)$$

gdzie:

- $N_k$  – końcowa liczba wykiełkowanych nasion
- $N(t)$  – liczba nasion wykiełkowanych po danym czasie  $t$
- $\alpha$  – współczynnik szybkości kiełkowania [1/h]
- $t_0$  – czas wykiełkowania pierwszego z wysianych nasion

Przykładowe dopasowanie krzywej logistycznej jest przedstawione na rys. 1 dla danych z pojedynczego poletka.



Rys. 1. Punkty otrzymane dla procesu kiełkowania jednej z replikacji ( $y$ ) i dopasowana do nich krzywa logistyczna  $N(t)$

Fig. 1. Points obtained for sprouting process of one from replications ( $y$ ) and logistic curve  $N(t)$  fitted to them

Jako kryterium sprawdzające dobroć dopasowania zastosowano często wykorzystywany do tego celu współczynnik determinacji  $R^2$  wyrażony wzorem [Ryan 1997, Seber i Wild 1989]:

$$R^2 = 1 - \frac{SSE}{SST} \quad (2)$$

gdzie:

$y_i$  – oznaczają zebrane dane doświadczalne,

$f_i$  – wartości dopasowywanej funkcji odpowiadające wartościom  $y_i$

zaś

$$SSE = \sum_{i=1}^n (y_i - f_i)^2 \quad (3)$$

$$SST = \sum_{i=1}^n (y_i - \bar{y})^2$$

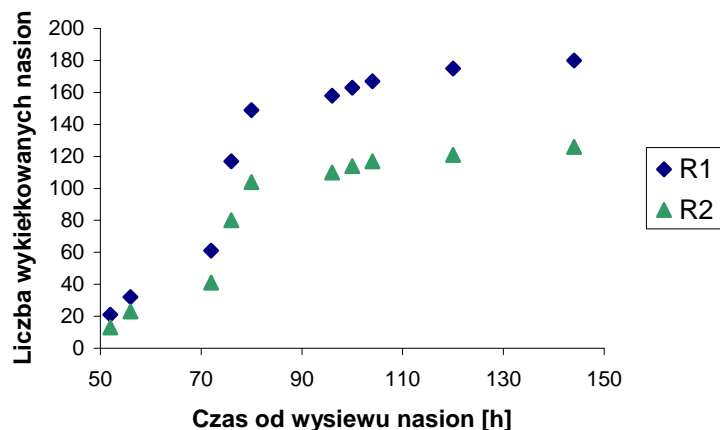
Wartości tego współczynnika bliskie 1 wskazują na dobre dopasowanie, natomiast bliskie 0 na słabe dopasowanie. Jeżeli badana funkcja całkowicie odbiega od danych doświadczalnych, wtedy ten współczynnik może osiągnąć wartość ujemną, co świadczy o tym, że trzeba taką funkcję odrzucić, bo nie jest ona właściwym modelem opisującym przebieg procesu. W ogólnym przypadku krzywą logistyczną można opisać wzorem z trzema parametrami:

$$f(t) = \frac{A}{1 + B e^{Ct}} \quad (4)$$

Analiza problemu poruszonego w pracy dotyczy jedynie parametru A, gdyż jest on najsilniej związany ze zdolnością kiełkowania i ma zasadnicze znaczenie przy próbie oszacowania wspólnej krzywej dla kilku replikacji.

### Problem dopasowania wspólnej krzywej

W praktyce najczęściej mamy do czynienia z sytuacją, gdy dla badanej odmiany zastosowane są powtórzenia, mające na celu uzyskanie bardziej precyzyjnych oszacowań. Różne czynniki, takie jak zróżnicowane warunki glebowe, pogodowe czy różnego rodzaju zdarzenia losowe, mogą wpłynąć w znaczący sposób na wyniki otrzymane dla poszczególnych replikacji. Taka sytuacja przedstawiona jest na wykresie 2, na którym zobrazowana jest liczba wykiełkowanych ziaren dla dwóch replikacji formy mieszaniec „Presto” x PPG 115.

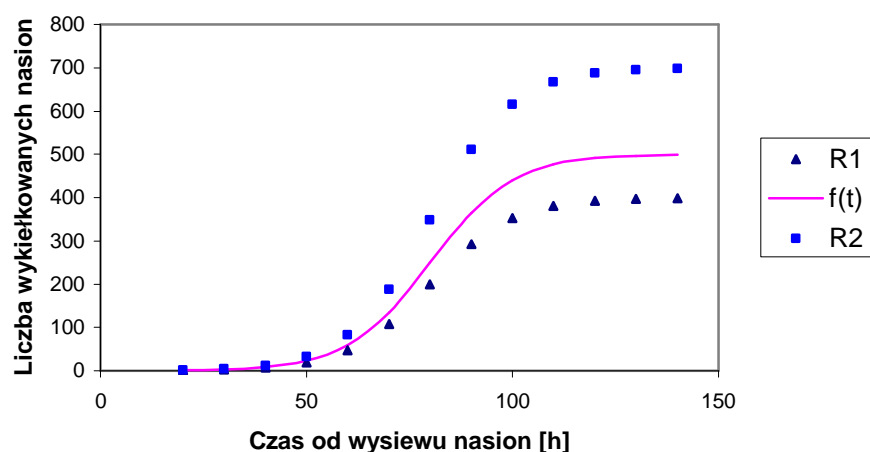


Rys. 2. Punkty odzwierciedlające proces kiełkowania przy dwóch replikacjach R1 i R2

Fig. 2. Points representing sprouting process for two replications R1 and R2

W sytuacji, gdy uwzględniono kilka replikacji celowe jest dopasowanie wspólnej krzywej, która mogłaby zobrazować proces kiełkowania dla wszystkich replikacji. Niekiedy dane otrzymane w poszczególnych replikacjach różnią się na tyle, iż zdaje się, że niemożliwe jest dopasowanie takiej krzywej. Powstającym problemem jest więc sposób jej oszacowania. Został on dla uproszczenia sytuacji rozpatrzony na przykładzie dwóch replikacji. Przede wszystkim podjęto próbę określenia, kiedy jest możliwy dobór takiej wspólnej krzywej.

W wyniku podjętych analiz dla krzywej logistycznej, którą można oszacować proces kiełkowania ziaren zbóż stymulowanych polem magnetycznym zaobserwowano następujący fakt przedstawiony na wykresie 3.



Rys 3. Wykres krzywej  $f(t)$  o takim samym dopasowaniu do danych dla dwóch replikacji R1 i R2 według współczynnika  $R^2$

Fig. 3. Graph of curve  $f(t)$ , which is fitted in the same degree to replications R1 and R2 according to coefficient  $R^2$

Wykres przedstawia krzywą logistyczną oznaczoną przy pomocy ciągłej linii oraz obserwacje dla dwóch teoretycznych replikacji, z których punkty R1 położone są nad rozważaną krzywą, a punkty R2 pod nią. Obserwacje te stanowią punkty leżące na krzywych logistycznych o innym parametrze niż opisana wcześniej krzywa. Jak się okazuje dopasowanie rozpatrywanej krzywej  $f(t)$  zarówno do punktów replikacji R1 jak i R2 charakteryzuje się taką samą wartością współczynnika  $R^2$ . Wynika z tego, że rezultaty uzyskane dla replikacji o słabszej zdolności kiełkowania, czyli gdy punkty leżą poniżej krzywej, są bardziej narażone na niższą wartość

współczynnika, a więc słabsze dopasowanie. Z tego względu można by pokusić się o następujące metody doboru krzywej, która byłaby dobrym dopasowaniem zarówno dla górnej, jak i dolnej replikacji: dopasowanie krzywej do wszystkich punktów zarówno z pierwszej, jak i z drugiej replikacji, dopasowanie do średniej arytmetycznej z obu replikacji, dopasowanie do średniej ważonej, uwzględniającej fakt, że punkty dla replikacji charakteryzującej się mniejszą zdolnością kiełkowania, muszą leżeć bliżej poszukiwanej krzywej. Dwie pierwsze metody dają ten sam wynik. Sprawą do rozstrzygnięcia pozostaje ocena dopasowania do średniej ważonej. Podjęto także próbę oceny, na ile odległe mogą być replikacje, aby można było dopasować wspólną krzywą.

### Wyniki badań

W tabeli 1 przedstawiono wartości współczynników  $a$  i  $b$  dla średniej ważonej:

$$\acute{s}r_w = a y_{R1} + b y_{R2}, \quad a+b=1 \quad (5)$$

gdzie

- $y_{R1}$  – obserwacja z replikacji R1
- $y_{R2}$  – obserwacja z replikacji R2

Wartości współczynników  $a$  i  $b$  wyznaczone zostały w taki sposób, aby wspólna krzywa miała taki sam współczynnik dopasowania do obserwacji otrzymanych dla obu replikacji. Oceny wartości współczynników dla średniej ważonej  $\acute{s}r_w$  dokonano dla pomiaru liczby wykiełkowanych nasion dokonanego w ostatnim punkcie czasowym, czyli w punkcie wyznaczającym dla poszczególnych replikacji osiągnięte przez nie ostatecznie zdolności kiełkowania. W rozważaniach przyjęto, że replikacja R1 charakteryzuje się mniejszą zdolnością kiełkowania niż replikacja R2.

Tabela 1. Współczynniki średniej ważonej

Table 1. Coefficients of weighted average

$R^2$	0	0,25	0,5	0,6	0,75	0,8	0,9	0,95	0,98
$a$	0,82	0,78	0,73	0,70	0,66	0,64	0,60	0,57	0,46
$b$	0,18	0,22	0,27	0,30	0,34	0,36	0,40	0,43	0,54

Tabela 2 przedstawia oszacowania odchyłeń wykresów dla replikacji R1 i R2 od krzywej tak, aby punkty dla poszczególnych replikacji miały takie same współczynniki  $R^2$ . Obliczenia przeprowadzono dla kilku ustalonych wartości współczynnika  $R^2$ . Dodatkowo w tabeli uwzględniono różnicę względną (w stosunku do

replikacji o większej liczbie wykiełkowanych nasion) między liczbą ziaren  $n_{R1}$  i  $n_{R2}$  wykiełkowanych w poszczególnych punktach czasowych dla obu replikacji i stosunek, mówiący jaką część liczby  $n_{R2}$  wykiełkowanych nasion dla replikacji o większej zdolności kiełkowania stanowi liczba  $n_{R1}$  nasion wykiełkowanych dla replikacji o mniejszej zdolności kiełkowania.

Tabela 2. Odchylenia replikacji R1 i R2 od średniej dające tę samą wartość współczynnika  $R^2$

Table 2. Deviations of replications R1 and R2 from average giving the same value of coefficient  $R^2$

$R^2$	0	0,25	0,5	0,6	0,75	0,8	0,9	0,95	0,98
Odchylenie R2	177,9%	124,4%	82,7%	68,0%	47,0%	40,1%	25,3%	16,7%	9,9%
Odchylenie R1	39,0%	35,6%	31,1%	28,8%	24,2%	22,2%	16,8%	12,5%	8,3%
$(n_{R2}-n_{R1})/n_{R2}$	0,780	0,713	0,623	0,576	0,484	0,445	0,336	0,250	0,166
$n_{R1}/n_{R2}$	0,220	0,287	0,377	0,424	0,5166	0,555	0,664	0,750	0,834

Tabela 3 przedstawia oszacowanie wartości współczynnika  $R^2$  oceniającego dopasowanie wspólnej krzywej do poszczególnych replikacji. Rozważana wspólna funkcja stanowi najlepsze dopasowanie dla średniej arytmetycznej w poszczególnych punktach pomiarowych i jest równoodległa w sensie zwykłej odległości od punktów obu replikacji. Rozważania przeprowadzono dla kilku przykładowych wielkości odchylenia replikacji od wspólnej krzywej. Także w tym przypadku oszacowano, podobnie jak w tabeli 2, różnicę względną i stosunek między liczbami wykiełkowanych nasion  $n_{R1}$  i  $n_{R2}$  dla obu replikacji.

Tabela 3. Wartości współczynnika  $R^2$  przy replikacjach o takim samym odchyleniu od średniej

Table 3. Values of coefficient  $R^2$  with replications equally deviated from average

Odchylenie od krzywej	5%	10%	15%	20%	25%	30%	35%	40%
$R^2$ przy replikacji R2	0,994	0,980	0,958	0,932	0,902	0,870	0,836	0,801
$R^2$ przy replikacji R1	0,993	0,970	0,924	0,848	0,729	0,552	0,293	-0,084
$(n_{R2}-n_{R1})/n_{R2}$	0,095	0,182	0,261	0,333	0,400	0,462	0,519	0,571
$n_{R1}/n_{R2}$	0,905	0,818	0,739	0,667	0,600	0,538	0,481	0,429

Parametry krzywej logistycznej, dla której prezentowane są wyniki były następujące  $A=500$ ,  $B=3000$ ,  $C=-0,11$ , ale sytuacja jest identyczna także dla innych wartości z zastrzeżeniem, że badane były tylko zmiany wartości parametru  $A$ .

## Wnioski

Podsumowując rozważania zawarte w pracy można wysunąć następujące wnioski:

1. Odległość punktów charakteryzujących proces kiełkowania od wspólnej krzywej ma różny wpływ na wartość współczynnika  $R^2$  dla poszczególnych replikacji w zależności od tego, czy reprezentują one replikacje o niższej czy wyższej zdolności kiełkowania (tabela 3).
2. Funkcja, która jest najlepiej dopasowana do średnich obliczonych dla poszczególnych punktów czasowych nie jest według kryterium  $R^2$  jednakowo dobrze dopasowana do obserwacji pochodzących tylko z pojedynczych replikacji. Gorsze dopasowanie zaobserwowano dla replikacji o słabszej zdolności kiełkowania.
3. Można stosować współczynnik  $R^2$  w celu określenia dopasowania krzywej do danych przy różnicy względnej liczby wykiełkowanych roślin nie przekraczającej wartości 0,25. Odpowiada to odchyleniu obserwacji ze słabiej kiełkującej replikacji od średniej arytmetycznej o ok. 15% (tabela 3). Natomiast całkowicie niewłaściwe jest stosowanie tego kryterium przy wspomnianej różnicy względnej powyżej wartości 0,55. Wtedy wartości współczynnika  $R^2$  osiągają wartości ujemne, świadczące o braku jakiegokolwiek dopasowania.
4. Pokazano, że w celu uzyskania krzywej logistycznej o jednakowym współczynniku  $R^2$  określającym dopasowanie do punktów dla każdej replikacji z osobna jest zastosowanie średniej ważonej z odpowiednio dobranymi wagami (tabela 1). Jednak zastosowanie takiej średniej ma charakter bardziej teoretyczny niż praktyczny.
5. Stwierdzono, że powszechnie stosowany współczynnik  $R^2$  nie jest najlepszym kryterium oceny dobroci dopasowania krzywych do punktów otrzymanych z doświadczenia, szczególnie wtedy, gdy wartości obserwacji z poszczególnych replikacji znacząco się od siebie różnią. Słuszne jest poszukiwanie innego lepszego kryterium.

## Bibliografia

- Pietruszewski S. 1998. Stanowisko do przedsięwziętej biostymulacji nasion zmiennym polem magnetycznym. *Inżynieria Rolnicza* 2(3): 31-36
- Pietruszewski S., Kornarzyński K. 1999. Wpływ pól magnetycznych na proces kiełkowania nasion. *Inżynieria Rolnicza* 2(8): 13-20
- Pietruszewski S., Kornarzyński K. 2001. Kiełkowanie nasion pszenicy jarej odmiany Jasna w stałym polu magnetycznym. *Fragmenta Agronomica* (XVIII), nr 2 (70).



Ryan T.P. 1997. Modern Regression Methods, Wiley & Sons, New York

Seber G. A. F., Wild. C. J. 1989. Nonlinear regression, New York: John Wiley & Sons

**A PROBLEM OF FITTING A COMMON CURVE  
FOR TWO REPLICATIONS  
ON EXAMPLE OF SPROUTING PROCESS  
OF CORN SEEDS STIMULATED WITH MAGNETIC FIELD**

**Summary**

A problem of fitting a common curve representing examined process, when there are two replications, is presented in the paper. The chosen curve should be good fitted for both replications. Logistic curve, which can be used to estimate sprouting process of corn seeds stimulated with magnetic field, was chosen to illustrate this problem. A determinant coefficient  $R^2$  was used as a criterion of goodness of fit. A method of initial estimation, when such a common curve can be fitted, is given. This paper presents different methods of choosing a common curve with coefficient  $R^2$  as a criterion and their observed properties.

**Key words:** curve fitting, logistic curve, coefficient  $R^2$ , magnetic field stimulation